Systematic Prediction of the Products and Intermediates of Isotopic Labeling in Reaction Pathway Studies

ANDREW V. ZEIGARNIK*, RAÚL E. VALDÉS-PÉREZ

Compuer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

Received 30 June 1997; accepted 17 December 1997

ABSTRACT: Isotopic labeling experiments can be highly informative in reaction pathway studies, but inferring the implications of a mechanistic hypothesis can be difficult, especially in the case of complex reactions. We report systematic methods for predicting the distribution of labeled products and intermediates given: (1) a mechanistic hypothesis; and (2) a proposed labeling experiment. The methods have been implemented with MECHEM—a computer aid for elucidating reaction mechanisms. As an illustration, we predict the outcomes of ethylene and propylene hydrogenation and *n*-heptane dehydrocyclization, for a variety of mechanisms and labeling experiments.
 J Comput Chem 19: 741–753, 1998

Keywords: isotopic labeling; reaction mechanism; elucidation; MECHEM program

* *Permanent address:* Laboratory of Chemical Kinetics and Catalysis Lomonosov Academy of Fine Chemical Technology, Moscow 117571, Russia.

Correspondence to: R. E. Valdés-Pérez; e-mail: valdes@cs.cmu.edu.

Contract/grant sponsor: National Science Foundation; contract/grant number: IRI-9421656.

Contract/grant sponsor: NSF Division of International Programs.

 CCC 0192-8651 / 98 / 070741-13

Introduction

Isotopic labeling experiments are frequently used in mechanistic studies of chemical reactions because isotopes give evidence on the intermediate species and elementary steps. The mechanisms of many reactions, especially in catalysis, were proven by isotopic labeling techniques.[1,2] The main types of isotopic techniques used in mechanistic studies of chemical reactions are: (i) kinetic isotope effect studies; and (ii) tracing the position of isotopic labels in intermediates and products. This article concerns only the second method.

Labeled reactants are always more costly than their unlabeled analogues. Therefore, it is useful to figure the results of labeling before proceeding to experiments. However, it is not always trivial to predict all possible positions of isotopic labels in products and intermediates and to know in advance whether an isotopic labeling experiment will be informative. This is especially true when the mechanism is long.

The aim of this study is to describe algorithms and an implemented computer program that can predict systematically the products and intermediates of a given labeling experiment on a given reaction mechanism. To our knowledge, no computer program of this sort has ever been reported, although the advantages are clear: the program avoids errors; the list of possible labeled intermediates is complete; and the search for possible distribution of labels takes, in many cases, a few minutes. The goal of our study was to design such a program and explore its application.

Problem Statement

Our understanding of the problem is as follows. Isotope tracing studies can be classified according to the number of types of possible labels (D, ^{13}C, ^{18}O, etc.). In chemistry, studies using a single label are most abundant, partly because of the reagent cost and partly because predicting the results of multiple-label experiments is even more difficult than in ordinary experiments.[3] Also, there may be qualitative and quantitative isotope tracing experiments. Qualitative experiments are based on a ''yes/no'' type of reasoning. In this case, on predicts whether or not the intermediate or product should be present. Quantitative experiments involve calculating the steady-state (or non-steady-state) ratios among the labeled products. Steady-state quantitative estimations of the product ratios are based on algebraic calculations. The calculation of the non-steady-state product ratios requires knowledge of the values of rate constants and concentrations of starting materials. This type of calculation is based on the numerical solution of sets of differential equations.

This article addresses the qualitative, single-label prediction problem. Extensions to multiple labels and to deriving quantitative information are future steps. The program made for this purpose accepts as input the mechanism and the list of starting materials containing the label. The program output is a list of attainable products and intermediates and a second list of active elementary steps involving the labeled species. This computer program was implemented as a component within the larger system, MECHEM, which was developed during the last few years.[4]

Algorithm

The algorithm for the qualitative prediction of single-label products and intermediates assumes the following as starting point:

- A reaction mechanism.
- A target element to be labeled (e.g., H) and a label (D).
- A set of labeled starting materials, which hereafter is referred to as a computer labeling experiment, or simply *experiment*.
- A specification of which steps of the mechanism will be treated as reversible.
- Any other steps to be included by the user (e.g., degenerate steps).

We distinguish between the mechanism's steps and user-added steps because of the context of our work, in which a computer program typically generates candidate mechanisms, but does not (currently) propose degenerate steps. The subsequent algorithm treats uniformly these steps of different origin.[5] Also, when a step is declared reversible, its reverse orientation is explicitly added to the list.

Our first design was a brute-force method that consisted of two basic stages: (i) explode every step into all possible labeled variants; and (ii) simulate the aggregate network, consisting of the original steps and all labeled variants, by running

the network on the labeled starting materials, thus obtaining all reachable products and intermediates, along with a record of every step that was activated. However, this method proved terribly slow for the *n*-heptane dehydrocyclization reaction discussed in what follows, so it was abandoned in favor of the following.

The improved algorithm consists of five stages:

(i) Get a reactant or reactant pair (call it or them R) that has not yet been ''run'' through the original reaction network; at first, R can only be drawn from the starting materials of the labeling experiment; if no waiting reactants remain, then exit by reporting the recorded steps and recorded products, both labeled and unlabeled.

(ii) Match R against the reactants of every network step, but ignore the presence of isotopes in R while doing the match.

(iii) For every successfully matched step, expand the step's products; that is, form all possible variants that have the same number of labels as in R, while pruning any detectable symmetric variants; each matched step then results in a variety of labeled steps that we will call the expanded network.

(iv) Compute the chemical distance[6] of every labeled step in the expanded network and remove all steps having a distance greater than (it cannot be less than) the chemical distance of the original, unlabeled step.

(v) Record the surviving steps, record the various labeled products, and add the latter to the ''pot'' from which waiting reactant pairs are drawn; continue by re-entering stage (i).

The precise algorithm design and programming for the aforementioned stages are relatively straightforward, except possiby for ''expand the step's products'' in stage (iii), which we proceed to describe in more detail by making use of an example.

PRUNING SYMMETRIC LABELINGS

To be specific, we will consider that, in the labeling experiment, some hydrogen atoms are labeled with deuterium (D); carbon and oxygen labelings are handled similarly. We also illustrate the procedures with reference to mechanism 2 from the propylene hydrogenation example in what follows, and when individual steps are needed for illustration, the fourth step from that mechanism.

$$MH + MCH(CH_3)_2 \rightarrow 2M + CH_3CH_2CH_3$$

In this article, M always denotes a metal site on the catalyst surface. For every product of the given step, place its H atoms into the set S. The products $2M + CH_3CH_2CH_3$, for example, lead to an S having eight H atoms. If, say, there are two D labels in R (which was matched against the step's reactants), then form all 28 ($= 8!/[8 - 2)!2!]$) ways to select (and thus label) two H atoms from the available eight. However, many of the 28 alternative product labelings are redundant for reasons of symmetry. (For example, all double H labels in a terminal carbon of propane are equivalent.) to detect and remove these redundancies, we partition S into subsets Y_i whose members play symmetric roles in the products. Thus, if we rewrite the products $2M + CH_3CH_2CH_3$ as $2M + CH^1H^2H^3CH^4H^5CH^6H^7H^8$ to distinguish the H atoms, the subsets Y_i become $(H^1H^2H^3)$, (H^4H^5), and $(H^6H^7H^8)$; that is, there are three sets of internally equivalent H atoms, each of which is useful for symmetry testing because it contains more then one atom.[7] Then, the 28 ways to select double H labels are pruned by removing redundant alternatives as determined by these Y_i values, resulting in a set of only six distinct labelings. (Note that this set has subtler symnmetries not currently removed, e.g., arising from the two equivalent methyl groups attached to the central carbon.)

The pruning procedure, which is short but somewhat complicated, can be described as follows:

```
Prune a set L of labeled atoms if this
test fails for some equivalence set Yi:

INITIALIZE
start? = true
last-atom-was-present?=false

ITERATE for each atom in Yi
this-atom-is-present? = true if the
atom is in L otherwise false
one of these propositions must be
true, otherwise exit with failure:
  1. start?
  2. last-atom-was-present?
```

```
  3. not(this-atom-is-present?);
last-atom-was-present? = this-atom-
is-present?
start? = false
```

For example, the set $L = (D^1D^3)$ of labeled atoms, arising from the labeling

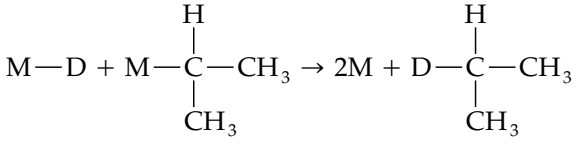
$$CD^1H^2D^3CH^4H^5CH^6H^7H^8,$$

would be rejected on the set $Y_i = (H^1H^2H^3)$ at the final iteration. (D^1D^3) is redundant because the labeling $CD^1D^2H^3CH^4H^5CH^6H^7H^8$ is the preferred symmetric variant. This completes the detailed description of ''expand the step's products'' in stage (iii).

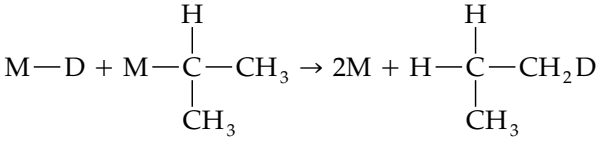
CHEMICAL DISTANCE

The use of chemical distance in stage (iv) to ensure that the labeled step is chemically identical to the original step may not be fully correct, but it is an excellent criterion for practical purposes. Alternatively, one might use a complex atom-by-atom and bond-by-bond matching to ensure that the initial and expanded steps reflect identical chemistries. Clearly, the use of chemical distance does not reject correct matches, and the risk that some false matches will be retained is minor in our opinion. Here are examples of correct and false matches for the step considered earlier:

Correct match:



False match:



The false match is intercepted by the chemical-distance criterion, because it implies five changes to bonding, rather than the three changes of the original step and of the correct match.

OUTPUT

Upon exit of the algorithm, all successful steps are reported and the labeled products and intermediates are clustered into groups based on a common, unlabeled origin. Currently, the MECHEM program incorporates the above algorithm and user interaction; the applications that are explored in the remainder of this work were done using MECHEM.

IMPORTANCE OF DEGENERATE STEPS

When analysing the position of labels in the products and intermediates, it is important to include degenerate steps in the mechanism if they do occur. Degenerate steps are those having identical species on the right-hand and left-hand sides of the chemical equation of a step. For example, the reaction:

$$ROOH + {}^{\bullet}OOR \leftrightarrows ROO^{\bullet} + HOOR$$

is degenerate if R is the same in both ROOH and $^{\bullet}$OOR. Degenerate steps may not affect the rate law of the reaction, but their inclusion into the mechanism may alter the pattern of isotopic label positions in intermediates and products. Also, one of the species may be involved on both the left and right sides, whereas the other species is different[8]:

$$Cl^- + CH_2{=}C{=}CHCH_2Cl \rightarrow CH_2{=}C(Cl)CH{=}CH_2 + Cl^-$$

Here the chlorine anion attacks one carbon atom and the other chlorine anion abstracts from another carbon atom.

MECHEM describes chemical structures as ordinary graphs. Therefore, a species with delocalized bonds like in η^3-allyl should be described as graphs (e.g., σ-allyl). However, when predicting the results of isotopic labeling, one should take into account that hydrogens adjacent to the terminal carbons in η^3-allyl are indistinguishable. Therefore, it is necessary to include degenerate steps like the following (carbon atoms are numbered for clarity):

$$M{-}C^1H_2{-}C^2H{=}C^3H_2 \leftrightarrows C^1H_2{=}C^2H{-}C^3H_2{-}M$$

Currently, identifying possible degenerate steps and adding them to the mechanism are the responsibility of the user.

Horiuti–Polanyi Mechanism for Propylene Hydrogenation

As a test example we used the mechanisms for the hydrogenation of propylene on Pt(111) discussed in the recent article by Cremer et al.[9] The possibility of propylidyne species formation on the surface was neglected for simplicity of illustration. Hence, these hypothetical mechanisms refer to the conditions considered in that study. However, the method in general is independent of the experimental conditions. Below are three mechanisms examined with the program. The first two pathways differ only by which carbon bonds to the metal within the last precursor of propane: the central carbon in one case, the end carbon in the other.

Mechanism 1:

1. $H_2 + 2M \rightleftarrows 2MH$
2. $C_3H_6 + 2M \rightleftarrows MCH(CH_3)—CH_2M$
3. $MH + MCH(CH_3)—CH_2M \rightleftarrows 2M + MCH_2—CH_2—CH_3$
4. $MH + MCH_2—CH_2—CH_3 \rightarrow 2M + CH_3—CH_2—CH_3$

Mechanism 2:

1. $H_2 + 2M \rightleftarrows 2MH$
2. $C_3H_6 + 2M \rightleftarrows MCH(CH_3)—CH_2M$
3. $MH + MCH(CH_3)—CH_2M \rightleftarrows 2M + M—CH(CH_3)_2$
4. $MH + M—CH(CH_3)_2 \rightarrow 2M + CH_3—CH_2—CH_3$

Mechanism 1 + 2:

1. $H_2 + 2M \rightleftarrows 2(MH)$
2. $C_3H_6 + 2M \rightleftarrows MCH(CH_3)—CH_2M$
3. $MH + MCH(CH_3)—CH_2M \rightleftarrows 2M + MCH(CH_3)_2$
4. $MH + MCH(CH_3)—CH_2M \rightleftarrows 2M + MCH_2—CH_2—CH_3$
5. $MH + MCH(CH_3)_2 \rightarrow 2M + CH_3—CH_2—CH_3$
6. $MH + MCH_2—CH_2—CH_3 \rightarrow 2M + CH_3—CH_2—CH_3$

The computer labeling experiments were $CH_2CHCH_3 + D_2$ and $CH_2CHCD_3 + H_2$. Note that these experiments cannot answer the question of whether the intermediate species is di-σ-bonded or π-bonded propylene. Therefore, we neglected this difference.

All steps except the final step of hydrogenation of intermediate species, which is commonly considered rate-limiting,[9] were treated as reversible. We neglected H/D exchange reactions of propane and propylene that in principle may occur over platinum single crystal surfaces,[10] because otherwise, all conceivable distributions of deuterium labels in propane will be possible and the example would not serve the illustration. Moreover, deuterium exchange in propane can be ruled out at low temperatures.

Continuing with the example, we assume that the available hypotheses correspond to the two primary mechanisms and to a third hypothesis that includes all of the steps from the two primary mechanisms.

The results of the computer experiments are summarized in Appendix A, from which rules for further experiment design can be inferred. Examples of such rules are as follows:

1. Experiment $CH_2CHCD_3 + H_2$ tends to give more information than experiment $CH_2CHCH_3 + D_2$, because the difference between the predicted products and intermediates is more pronounced. However, there is no strict dominance: the latter experiment can lead to crucial products that are not discriminated by the former experiment. In the case of experiment $CH_2CHCH_3 + D_2$, 11 species should appear no matter whether mechanism 1 or 2 is true, whereas, in $CH_2CHCD_3 + H_2$, this ''overlap'' is only five species.
2. d_8-propane, $CD_2HCDHCD_2H$, $CDH_2CD_2CD_3$, $CH_3CD_2CD_3$, $CDH_2CD_2CD_2H$, CH_3CDHCD_3, $CH_3CD_2CD_2H$, and $CDH_2CD_2CDH_2$ are not formed unless mechanism 1 + 2 is true.
3. In the case of mechanism 1 + 2, all possible deuterated and semideuterated products and intermediates are formed. Also, the spectrum of products is more diverse for mechanism 2 than for mechanism 1.
4. If nondeuterated propane is observed in the experiment $CH_2CHCD_3 + H_2$, then mechanism 2 is true.

5. If experiments $CH_2CHCD_3 + H_2$ and $CH_2CHCH_3 + D_2$ give the same set of deuterated products, then mechanism 2 is true.
6. If deuterium atom is bonded to carbon in position 2 of propane in experiment $CH_2CHCD_3 + H_2$, than mechanism 1 is wrong.
7. d_8—d_4-propanes are never observed if mechanism 1 is true.

This list is by no means complete: more rules could be inferred from analysis of the data on intermediates and starting materials. (One could also conceive of a separate computer program that generates these rules.) Using these rules and analyzing the tabulated predictions, one may compare the results of labeling experiments and choose the most informative one (or several of the most informative experiments).

As can be seen from the predictions in Table I, the number of ''expanded'' elementary steps may be very large, and it is sometimes difficult to write them by hand. This is a further argument in favor of the use of automatic formal methods. The expanded elementary steps for mechanism 1 and experiment $C_3H_6 + D_2$ are listed in Appendix B.

Ethylene Hydrogenation

When predicting the distribution of isotopic labels, it is often useful to generate systematically the mechanisms of reactions and then examine them from the standpoint of isotopic labeling. Proceeding thus, one is likely to overlook some mechanistic possibilities that also account for any specific experimental distribution of isotopic labels. Moreover, the ''space'' of examined mechanisms will be known and thus the framework in which isotopic labeling experiments give mechanistic evidence is better characterized.

TABLE I.
Number of Labeled and Unlabeled Steps in Reaction Network.

	$C_3H_6 + D_2$	$CH_2CHCD_3 + H_2$
Mechanism 1	24	7
Mechanism 2	98	98
Mechanism 1 + 2	358	358

We applied the computer program MECHEM[4] to generate the mechanisms of ethylene hydrogenation on the Pt(111) single-crystal surface covered with the ethylidyne species ($M_3C—CH_3$).[11] We generated the simplest mechanisms (fewest species and steps) using the minimal set of constraints given in what follows. More recent studies by Somorjai and coworkers,[12] as well as studies by other researchers, showed that ethylidyne species are spectators during the catalytic reaction. Nevertheless, this mechanism is suitable for the purpose of illustration, although it cannot be recommended for kinetic modeling.

The reactants which we input to the program were H_2, ethylene, and a pair of adjacent catalyst sites (single site M was not given to the program). The defined product was ethane. The list of constraints is:

- Every conjectured species (intermediate) has at most two carbon atoms.
- Every intermediate is bound to at most three metal atoms of the surface.
- The catalyst mediates all steps.
- No reactions between gas-phase and surface species.
- Reject the step $MCH_2—CH_3 \rightarrow CH_3—CH_3 + MCH_2—CH_2M$.
- $M_3C—CH_3$ is a necessary precursor of $CH_3—CH_3$.
- The overall reaction stoichiometry is $H_2 + CH_2{=}CH_2 \rightarrow CH_3—CH_3$.

Under these conditions, MECHEM generated three mechanisms in addition to the Zaera–Somorjai mechanism. Each of the four mechanisms contains the same intermediate species and the same number of steps. The steps are analogous across the mechanisms:

1. Adsorption of the ethylene on the surface (the nature of bonding is not considered here):

$$2M + CH_2{=}CH_2 \rightarrow MCH_2{-}CH_2M$$

2. Dissociative adsorption of H_2 on the metal surface:

$$H_2 + 2M \rightarrow 2MH$$

3. Initial hydrogenation of the adsorbed ethylene by MH:

$$MCH_2—CH_2M + MH \rightarrow 2M + MCH_2—CH_3$$

or by ethylidyne or ethylidene; for example, in the case of ethylidyne:

$$MCH_2—CH_2M + M_2CH—CH_3 \rightarrow M_3C—CH_3 + MCH_2—CH_3$$

4. Final hydrogenation by $M_2CH—CH_3$:

$$MCH_2—CH_3 + M_2CH—CH_3 \rightarrow M_3C—CH_3 + CH_3—CH_3$$

or by MH:

$$MCH_2CH_3 + MH \rightarrow 2M + CH_3—CH_3$$

5. Hydrogenation of ethylidyne to ethylidene by MH:

$$M_3C—CH_3 + MH \rightarrow 2M + M_2CH—CH_3.$$

or by $MCH_2—CH_3$:

$$M_3C—CH_3 + MCH_2—CH_3 \rightarrow MCH_2—CH_2M + M_2CH—CH_3$$

The mechanisms generated by computer are as follows:

Mechanism 1:

1. $2M + CH_2=CH_2 \rightarrow MCH_2—CH_2M$
2. $H_2 + 2M \rightarrow 2MH$
3. $MCH_2—CH_2M + MH \rightarrow 2M + MCH_2—CH_3$
4. $M_3C—CH_3 + MCH_2—CH_3 \rightarrow MCH_2—CH_2M + M_2CH—CH_3$
5. $MCH_2—CH_3 + M_2CH—CH_3 \rightarrow M_3C—CH_3 + CH_3—CH_3$

Mechanism 2:

1. $2M + CH_2=CH_2 \rightarrow MCH_2—CH_2M$
2. $H_2 + 2M \rightarrow 2MH$
3. $M_3C—CH_3 + MH \rightarrow 2M + M_2CH—CH_3$
4. $MCH_2—CH_2M + MH \rightarrow 2M + MCH_2—CH_3$
5. $M_2CH—CH_3 + MCH_2—CH_3 \rightarrow M_3C—CH_3 + CH_3—CH_3$

Mechanism 3 (Zaera–Somorjai mechanism):

1. $2M + CH_2=CH_2 \rightarrow MCH_2—CH_2M$
2. $H_2 + 2M \rightarrow 2MH$
3. $M_3C—CH_3 + MH \rightarrow 2M + M_2CH—CH_3$
4. $MCH_2—CH_2M + M_2CH—CH_3 \rightarrow M_3C—CH_3 + MCH_2—CH_3$
5. $M_2CH—CH_3 + MCH_2—CH_3 \rightarrow M_3C—CH_3 + CH_3—CH_3$

Mechanism 4:

1. $2M + CH_2=CH_2 \rightarrow MCH_2CH_2M$
2. $H_2 + 2M \rightarrow 2MH$
3. $M_3C—CH_3 + MH \rightarrow 2M + M_2CH—CH_3$
4. $MCH_2—CH_2M + M_2CH—CH_3 \rightarrow M_3C—CH_3 + MCH_2—CH_3$
5. $MH + MCH_2—CH_3 \rightarrow 2M + CH_3—CH_3$

We carried out computer experiments $C_2H_4 + D_2$ of two types: (1) consider all steps irreversible; and (2) consider ethylene adsorption and the initial hydrogenation of ethylene reversible (other steps are irreversible). These experiments showed that it would be inadvisable to trace isotopes among the intermediates, because the distributions of labels for all the mechanisms in each experiment are virtually the same. Tracing of isotopic labels in deuterated ethylenes also will not discriminate among the mechanisms. However, the distribution of labels in ethane appeared somewhat more sensitive to the mechanism and step reversibility.

If all steps are considered irreversible, then mechanisms 2, 3, and 4 form only the single deuterated product $CH_2D—CH_2D$, but mechanism 1 allows the whole spectrum of deuterated ethanes, although it does not yield nondeuterated ethane.

If ethylene adsorption and initial hydrogenation are treated as reversible steps, then mechanisms 1, 2, and 3 are indistinguishable: they give all conceivable ethanes, whereas mechanism 4 does not give nondeuterated ethane.

As can be seen, several very similar mechanisms may exist for such a simple reaction. Even in this case isotopic labeling may disprove some of them. However, there still may remain some indis-

tinguishable mechanisms that should be tested against other experimental evidence.

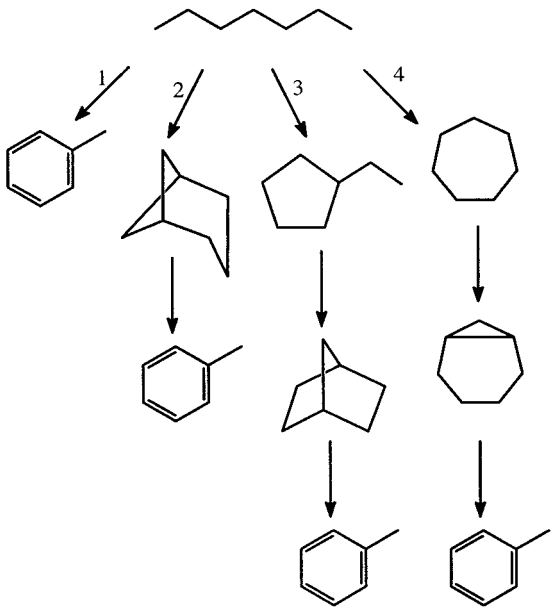
n-Heptane Dehydrocyclization

Our final example, taken from a book by Ozaki,[1] involves a polyvalent label (carbon), and hence demonstrates the method's generality. Pruning of symmetric structures in the case of polyvalent labels has not yet been worked out, so each run takes considerably longer (around 10–20 minutes). Osaki described early experiments on *n*-heptane dehydrocyclization. Mitchell found that the dehydrocyclization of [1-^{14}C]-*n*-heptane over $Cr_2O_3—Al_2O_3—K_2O$ gave only 27–29% methyl-labeled methylbenzene instead of 50% as predicted by the ''direct'' C_6 ring formation (the label is shown by a solid dot)[13]:

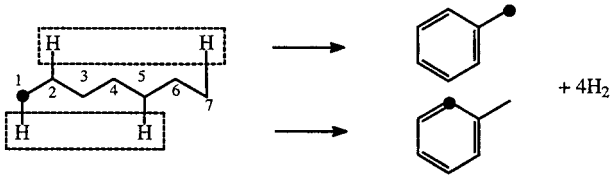
Then, Pines and Chen[14] repeated this experiment using various oxide catalysts of different acidity. Their results were also inconsistent with ''direct'' C_6 ring closure because the amount of methyl-labeled methylbenzene was also small. According to Ozaki,[1] this incompatibility led to various alternative suggestions, such as the formation of a bicyclic intermediate of bicyclo[3.1.1]cyclohexane structure[14]:

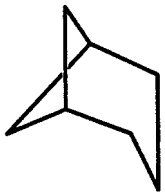
This hypothesis was examined by Feighan and Davis using [1-^{14}C]-*n*-heptane.[15]

We carried out computational experiments on the four possible positions of single ^{14}C labels and the four mechanisms shown in Scheme 1. Instead of dehydrocyclization to methylbenzene, we experimented with analogous saturated cyclic structures (which does not change the labeling pattern). Of course, these mechanisms only describe the skeletal rearrangements and thus neglect the role of a catalyst. Although more complex mechanistic hypotheses are within MECHEM's scope, our aim here is to apply the program to noncatalytic organic chemistry and polyvalent atomic labels.



SCHEME 1.

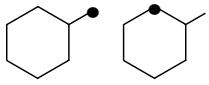
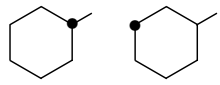
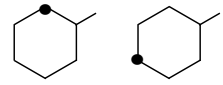
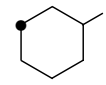
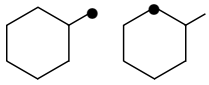
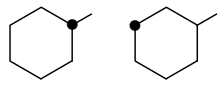
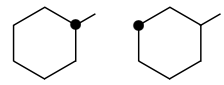
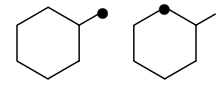
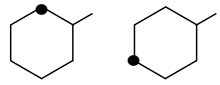
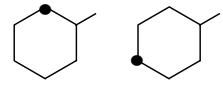
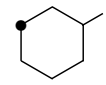




The results summarized in Table II show that two hypotheses (direct ''1,6-ring'' closure and the formation of bicyclic intermediate of bicyclo[3.1.1]cyclohexane) are indistinguishable by any single-^{14}C-label experiment. Therefore, Pines and Chen's suggestion[14] that this bicyclic intermediate is formed cannot be confirmed by experiments with [1-^{14}C]-*n*-heptane, [4-^{14}C]-*n*-heptane, or any other single carbon label. Other hypotheses proposed here for purposes of illustration can be discriminated. For example, starting with [1-^{14}C]-*n*-heptane, mechanism 4 can form all of the labeled isomers. Mechanism 3 forms more isomers than mechanisms 1 and 2 for any position of label in *n*-hexane.

It is worth mentioning that, before using the program, we tried to predict the distribution of labels using only pencil and paper. This took about half an hour, but the results did not completely agree with the program's subsequent predictions, which turned out—upon re-examination—to be correct.

TABLE II.
Predicted Positions of Isotopic Labels in Products of *n*-Heptane Dehydrocyclization.

Hypothetical mechanism (see Scheme 1)	Initial position of the label in *n*-heptane			
	[1-^{14}C]-*n*-heptane	[2-^{14}C]-*n*-heptane	[3-^{14}C]-*n*-heptane	[4-^{14}C]-*n*-heptane
1, 2				
3				
4	 All (five) positions	 All (five) positions	 All (five) positions	 All (five) positions

Solid circles stand for carbon labels.

Conclusion

We have formed the task of predicting the possible positions of isotopic labels in the products and intermediates of a given mechanism when some atoms in the starting materials are labeled with isotopes of a single type. The algorithm, which relies on detecting and suppressing some symmetric labelings, is implemented within MECHEM, an aid for elucidating reaction mechanisms. We have illustrated the method by predicting the discriminatory power of several possible labeling experiments, in the context of similar mechanisms for the catalytic hydrogenation of olefins and *n*-heptane dehydrocyclization. Future steps include improving the symmetry detection for both univalent and polyvalent labels, and addressing the single-isotope/quantitative and the multiple-isotope/qualitative prediction tasks. By the time this article is published we will have expanded the method's role in MECHEM so that it can serve as a constraint by, for instance, enabling a search for a mechanism that is (in)capable of forming a given labeled product. The authors invite inquiries from those wishing to apply these methods to problems of their own.

Appendix A

Products and Intermediates for Propylene Hydrogenation.

No.	Species	Mechanism 1 $C_3H_6 + D_2$	Mechanism 2 $C_3H_6 + D_2$	Mechanism 1 $CH_2CHCD_3 + H_2$	Mechanism 2 $CH_2CHCD_3 + H_2$	Mechanism 1 + 2 $C_3H_6 + D_2$
1	MD	Yes	Yes	No	Yes	Yes
2	MH	Yes	Yes	Yes	Yes	Yes
3	$MCD(CD_3)—CD_2M$	No	No	No	No	Yes
4	$MCD_2—CD_2—CD_3$	No	No	No	No	Yes
5	$MCD(CD_3)_2$	No	No	No	No	Yes
6	$MCH(CD_3)—CD_2M$	No	Yes	No	Yes	Yes
7	$MCD(CD_3)—CDHM$	No	No	No	No	Yes
8	$MCD(CD_2H)—CD_2M$	No	No	No	No	Yes
9	$MCD_2—CDH—CD_3$	No	No	No	No	Yes
10	$MCH(CD_3)_2$	No	Yes	No	Yes	Yes

(continued)

(Continued).

No.	Species	Mechanism 1 $C_3H_6 + D_2$	Mechanism 2 $C_3H_6 + D_2$	Mechanism 1 CH_2CHCD_3 $+H_2$	Mechanism 2 CH_2CHCD_3 $+H_2$	Mechanism 1 + 2 $C_3H_6 + D_2$
11	$MCDH-CD_2-CD_3$	No	No	No	No	Yes
12	$MCD_2-CD_2-CD_2H$	No	No	No	No	Yes
13	$MCD(CD_2H)(CD_3)$	No	No	No	No	Yes
14	$MCH(CD_3)-CDHM$	No	Yes	No	Yes	Yes
15	$MCD(CD_3)-CH_2M$	No	No	No	No	Yes
16	$MCH(CD_2H)-CD_2M$	No	Yes	No	Yes	Yes
17	$MCD(CD_2H)-CDHM$	No	No	No	No	Yes
18	$MCD(CDH_2)-CD_2M$	No	No	No	No	Yes
19	$MCD_2-CH_2-CD_3$	No	No	No	No	Yes
20	$MCDH-CDH-CD_3$	No	No	No	No	Yes
21	$MCH_2-CD_2-CD_3$	No	No	No	No	Yes
22	$MCD_2-CDH-CD_2H$	No	No	No	No	Yes
23	$MCH(CD_2H)(CD_3)$	No	Yes	No	Yes	Yes
24	$MCDH-CD_2-CD_2H$	No	No	No	No	Yes
25	$MCD(CD_2H)_2$	No	No	No	No	Yes
26	$MCD_2-CD_2-CDH_2$	No	No	No	No	Yes
27	$MCD(CDH_2)(CD_3)$	No	No	No	No	Yes
28	$MCH(CD_3)-CH_2M$	No	Yes	Yes	Yes	Yes
29	$MCH(CD_2H)-CDHM$	No	Yes	No	Yes	Yes
30	$MCH(CDH_2)-CD_2M$	No	Yes	No	Yes	Yes
31	$MCD(CD_2H)-CH_2M$	No	No	No	No	Yes
32	$MCD(CH_3)-CD_2M$	No	No	No	No	Yes
33	$MCD(CDH_2)-CDHM$	No	No	No	No	Yes
34	$MCDH-CH_2-CD_3$	No	No	No	No	Yes
35	$MCH_2-CDH-CD_3$	No	No	No	No	Yes
36	$MCD_2-CH_2-CD_2H$	no	No	No	No	Yes
37	$MCDH-CDH-CD_2H$	No	No	No	No	Yes
38	$MCH(CD_2H)(CD_2H)$	No	Yes	No	Yes	Yes
39	$MCD_2-CDH-CDH_2$	No	No	No	No	Yes
40	$MCH(CDH_2)(CD_3)$	No	Yes	No	Yes	Yes
41	$MCH_2-CD_2-CD_2H$	No	No	No	No	Yes
42	$MCD_2-CD_2-CH_3$	No	No	No	No	Yes
43	$MCD(CH_3)(CD_3)$	No	No	No	No	Yes
44	$MCDH-CD_2-CDH_2$	No	No	No	No	Yes
45	$MCD(CDH_2)(CD_2H)$	No	No	No	No	Yes
46	$MCH(CD_2H)-CH_2M$	No	Yes	No	Yes	Yes
47	$MCH(CH_3)-CD_2M$	No	Yes	No	Yes	Yes
48	$MCH(CDH_2)-CDHM$	No	Yes	No	Yes	Yes
49	$MCD(CDH_2)-CH_2M$	No	No	No	No	Yes
50	$MCD(CH_3)-CDHM$	No	No	No	No	Yes
51	$MCH_2-CH_2-CD_3$	No	No	Yes	Yes	Yes
52	$MCDH-CH_2-CD_2H$	No	No	No	No	Yes
53	$MCD_2-CH_2-CDH_2$	No	No	No	No	Yes
54	$MCH_2-CDH-CD_2H$	No	No	No	No	Yes
55	$MCD_2-CDH-CH_3$	No	No	No	No	Yes
56	$MCH(CH_3)(CD_3)$	No	No	No	No	Yes
57	$MCD-CDH-CDH_2$	No	No	No	No	Yes
58	$MCH(CDH_2)(CD_2H)$	No	Yes	No	Yes	Yes
59	$MCH_2-CD_2-CDH_2$	No	No	No	No	Yes
60	$MCD(CDH_2)_2$	No	No	No	No	Yes

(continued)

(Continued).

No.	Species	Mechanism 1 $C_3H_6 + D_2$	Mechanism 2 $C_3H_6 + D_2$	Mechanism 1 $CH_2CHCD_3 + H_2$	Mechanism 2 $CH_2CHCD_3 + H_2$	Mechanism 1 + 2 $C_3H_6 + D_2$
61	$MCDH-CD_2-CH_3$	No	No	No	No	Yes
62	$MCD(CH_3)(CD_2H)$	No	No	No	No	Yes
63	$MCH(CDH_2)-CH_2M$	No	Yes	No	Yes	Yes
64	$MCH(CH_3)-CDHM$	No	Yes	No	Yes	Yes
65	$MCD-CH_2M-CH_3$	Yes	No	No	No	Yes
66	$MCH_2-CH_2-CD_2H$	No	No	No	No	Yes
67	$MCD_2-CH_2-CH_3$	No	No	No	No	Yes
68	$MCDH-CH_2-CDH_2$	No	No	No	No	Yes
69	$MCH_2-CDH-CDH_2$	No	No	No	No	Yes
70	$MCH(CDH_2)_2$	No	Yes	No	Yes	Yes
71	$MCDH-CDH-CH_3$	No	No	No	No	Yes
72	$MCH(CH_3)(CD_2H)$	No	Yes	No	Yes	Yes
73	$MCH_2-CD_2-CH_3$	Yes	Yes	No	Yes	Yes
74	$MCD(CH_3)(CDH_2)$	No	No	No	No	Yes
75	$MCH(CH_3)-CH_2M$	Yes	Yes	No	Yes	Yes
76	$MCH_2-CH_2-CDH_2$	No	No	No	No	Yes
77	$MCDH-CH_2-CH_3$	No	No	No	No	Yes
78	$MCD(CH_3)_2$	No	No	No	No	Yes
79	$MCH_2-CDH-CH_3$	Yes	No	No	No	Yes
80	$MCH(CH_3)(CDH_2)$	No	Yes	No	Yes	Yes
81	$MCH_2-CH_2-CH_3$	Yes	No	No	No	Yes
82	$MCH(CH_3)_2$	No	Yes	No	Yes	Yes
83	D_2	Yes	Yes	No	Yes	Yes
84	HD	Yes	Yes	No	Yes	Yes
85	H_2	Yes	Yes	Yes	Yes	Yes
86	$CD_2=CD-CD_3$	No	No	No	No	Yes
87	$CD_3-CD_2-CD_3$	No	No	No	No	Yes
88	$CDH=CD-CD_3$	No	No	No	No	Yes
89	$CD_2=CH-CD_3$	No	Yes	No	Yes	Yes
90	$CD_2=CD-CD_2H$	No	No	No	No	Yes
91	$CD_3-CDH-CD_3$	No	Yes	No	Yes	Yes
92	$CD_2H-CD_2-CD_3$	No	No	No	No	Yes
93	$CH_2=CD-CD_3$	No	No	No	No	Yes
94	$CDH=CH-CD_3$	No	Yes	No	Yes	Yes
95	$CDH=CD-CD_2H$	No	No	No	No	Yes
96	$CD_2=CH-CD_2H$	No	Yes	No	Yes	Yes
97	$CD_2=CD-CDH_2$	No	No	No	No	Yes
98	$CD_3-CH_2-CD_3$	No	Yes	No	Yes	Yes
99	$CD_2H-CDH-CD_3$	No	Yes	No	Yes	Yes
100	$CD_2H-CD_2-CD_2H$	No	No	No	No	Yes
101	$CDH_2-CD_2-CD_3$	No	No	No	No	Yes
102	$CH_2=CH-CD_3$	No	Yes	Yes	Yes	Yes
103	$CH_2=CD-CD_2H$	No	No	No	No	Yes
104	$CDH=CH-CD_2H$	No	Yes	No	Yes	Yes
105	$CDH=CD-CDH_2$	No	No	No	No	Yes
106	$CD_2=CH-CDH_2$	No	Yes	No	Yes	Yes
107	$CD_2=CD-CH_3$	No	No	No	No	Yes
108	$CD_2H-CH_2-CD_3$	No	Yes	No	Yes	Yes
109	$CD_2H-CDH-CD_2H$	No	Yes	No	Yes	Yes
110	$CDH_2-CDH-CD_3$	No	Yes	No	Yes	Yes

(continued)

(Continued).

No.	Species	Mechanism 1 $C_3H_6 + D_2$	Mechanism 2 $C_3H_6 + D_2$	Mechanism 1 CH_2CHCD_3 $+H_2$	Mechanism 2 CH_2CHCD_3 $+H_2$	Mechanism 1 + 2 $C_3H_6 + D_2$
111	$CH_3—CD_2—CD_3$	No	No	No	No	Yes
112	$CDH_2—CD_2—CD_2H$	No	No	No	No	Yes
113	$CH_2=CH—CD_2H$	No	Yes	No	Yes	Yes
114	$CH_2=CD—CDH_2$	No	No	No	No	Yes
115	$CDH=CH—CDH_2$	No	Yes	No	Yes	Yes
116	$CDH=CD—CH_3$	No	No	No	No	Yes
117	$CD_2=CH—CH_3$	No	Yes	No	Yes	Yes
118	$CD_2H—CH_2—CD_2H$	No	Yes	No	Yes	Yes
119	$CDH_2—CH_2—CD_3$	No	Yes	No	Yes	Yes
120	$CH_3—CDH—CD_3$	No	No	No	No	Yes
121	$CDH_2—CDH—CD_2H$	No	Yes	No	Yes	Yes
122	$CDH_2—CD_2—CDH_2$	No	No	No	No	Yes
123	$CH_3—CD_2—CD_2H$	No	No	No	No	Yes
124	$CH_2=CH—CDH_2$	No	Yes	No	Yes	Yes
125	$CH_2=CD—CH_3$	Yes	No	No	No	Yes
126	$CDH=CH—CH_3$	No	Yes	No	Yes	Yes
127	$CH_3—CH_2—CD_3$	No	Yes	Yes	Yes	Yes
128	$CDH_2—CH_2—CD_2H$	No	Yes	No	Yes	Yes
129	$CDH_2—CDH—CDH_2$	No	Yes	No	Yes	Yes
130	$CH_3—CDH—CD_2H$	No	Yes	No	Yes	Yes
131	$CH_3—CD_2—CDH_2$	Yes	No	No	No	Yes
132	$CDH_2—CH_2—CDH_2$	No	Yes	No	Yes	Yes
133	$CH_3—CH_2—CD_2H$	No	Yes	No	Yes	Yes
134	$CH_3—CD_2—CH_3$	Yes	No	No	No	Yes
135	$CH_3—CDH—CDH_2$	Yes	Yes	No	Yes	Yes
136	$CH_3—CH_2—CDH_2$	Yes	Yes	No	Yes	Yes
137	$CH_3—CDH—CH_3$	Yes	Yes	No	Yes	Yes
138	$CH_3—CH_2—CH_3$	Yes	Yes	No	Yes	Yes

Appendix B

Reaction Network for Mechanism 1 of Propylene Hydrogenation and Experiment $C_3H_6 + D_2$.

1. $MD + MD \rightarrow 2M + D_2$
2. $2M + HD \rightarrow MD + MH$
3. $MD + MH \rightarrow 2M + HD$
4. $2M + CH_2=CD—CH_3 \rightarrow MCD—CH_2M—CH_3$
5. $MCD—CH_2M—CH_3 \rightarrow 2M + CH_2=CD—CH_3$
6. $2M + MCH_2—CD_2—CH_3 \rightarrow MD + MCD—CH_2M—CH_3$
7. $MD + MCH_2—CD_2—CH_3 \rightarrow 2M + CH_3—CD_2—CDH_2$
8. $MH + MCH_2—CD_2—CH_3 \rightarrow 2M + CH_3—CD_2—CH_3$
9. $MH + MCH_2—CDH—CH_3 \rightarrow 2M + CH_3—CDH—CH_3$
10. $MD + MCH_2—CH_2—CH_3 \rightarrow 2M + CH_3—CH_2—CDH_2$
11. $MD + MCD—CH_2M—CH_3 \rightarrow 2M + MCH_2—CD_2—CH_3$
12. $MH + MCD—CH_2M—CH_3 \rightarrow 2M + MCH_2—CDH—CH_3$
13. $H_2 + 2M \rightarrow 2MH$
14. $2MH \rightarrow H_2 + 2M$
15. $2M + MCH_2—CH_2—CH_3 \rightarrow MH + MCH(CH_3)—CH_2M$
16. $MH + MCH_2—CH_2—CH_3 \rightarrow 2M + CH_3—CH_2—CH_3$
17. $MH + MCH(CH_3)—CH_2M \rightarrow 2M + MCH_2—CH_2—CH_3$
18. $2M + MCH_2—CDH—CH_3 \rightarrow MH + MCD—CH_2M—CH_3$
19. $2M + MCH_2—CDH—CH_3 \rightarrow MD + MCH(CH_3)—CH_2M$
20. $MD + MCH_2—CDH—CH_3 \rightarrow 2M + CH_3—CDH—CDH_2$
21. $MD + MCH(CH_3)—CH_2M \rightarrow 2M + MCH_2—CDH—CH_3$
22. $2M + D_2 \rightarrow MD + MD$
23. $MCH(CH_3)—CH_2M \rightarrow CH_3—CH_2—CH_3 + 2M$
24. $CH_3—CH_2—CH_3 + 2M \rightarrow MCH(CH_3)—CH_2M$

References

1. A. Ozaki, *Isotopic Studies in Heterogeneous Catalysis*, Academic Press, New York, 1977.
2. S. Z. Roginskii, *Theoretical Principles of Isotope Methods for Investigating Chemical Reactions*, USSR Academy of Sciences, Moscow, 1956.
3. However, in biochemistry, studies in which several different labels are used simultaneously are becoming more frequent.
4. MECHEM has been written over the last 9 years and is currently a 60,000-line program, but it is under continual development. Except for the graphical user interface, it is entirely written in Lisp, including its molecules that implement algebraic and numerical constraints pertaining mainly to stoichiometry. The graphical interface, which is well advanced, is written in Tcl/Tk. Hence, the entire package will run on any platform that supports these two highly portable languages. More information about MECHEM and its availability is at www.cs.cmu.edu/~sci-disc/mechem.html. For references, see: (a) A. V. Zeigarnik, R. E. Valdés-Pérez, O. N. Temkin, L. G. Bruk, and S. I. Shalgunov, *Organometallics*, **16**, 3114 (1997); (b) R. E. Valdés-Pérez and A. V. Zeigarnik, *J. Mol. Catal. A: Chemical*, **119**, 405 (1997); (c) R. E. Valdés-Pérez, *Catal. Lett.*, **28**, 79 (1994); (d) R. E. Valdés-Pérez, *J. Comput. Chem.*, **13**, 1079 (1992); (e) R. E. Valdés-Pérez, *J. Comput. Chem.*, **15**, 1266 (1994); (f) R. E. Valdés-Pérez, *J. Chem. Inf. Comput. Sci.*, **34**, 976 (1994); (g) R. E. Valdés-Pérez, *J. Chem. Inf. Comput. Sci.*, **31**, 554 (1991); (h) R. E. Valdés-Pérez, *J. Comput. Chem.*, **14**, 1454 (1993); (i) L. G. Bruk, S. N. Gorodskii, A. V. Zeigarnik, R. E. Valdés-Pérez, and O. N. Temkin, *J. Mol. Catal. A: Chemical*, in press.
5. When a mechanism is input to MECHEM manually, the program does not recognize what really happens in each specific step and thus it will consider the transformation that follows the minimum structural change, which is not necessarily the intent of whoever devised the mechanism. The user can detect mismatches by asking the program for the chemical distance of each step and verifying that this distance matches the intent. If there is a mismatch, it may be possible to force the program to consider the more complex transformation by making use of constraints that prohibit the routes of fewer structural changes.
6. Chemical distance is defined here as the number of bonds that are broken or formed without account of bonds that change their multiplicity. A slightly different use of this concept can be found, for example, in: M. Wochner, J. Brandt, A. von Scholley, and I. Ugi, *Chimia*, **42**, 217 (1988).
7. Two monovalent atoms plat symmetric roles if they are bonded to the same atom, or to each other. Formalizing symmetry for polyvalent atoms is more difficult and has not yet been done. However, the number of polyvalent atoms is typically much fewer, so that the combinatorics is not as problematic as potentially would be the case for *H* atoms, say, without testing for symmetries.
8. O. N. Temkin, G. K. Shestakov, and Yu. A. Treger, *Acetylene: Chemistry, Reaction Mechanisms, and Technology*, Khimiya, Moscow, 1991 [in Russian].
9. P. S. Cremer, X. Su, Y. R. Shen, and G. A. Somorjai, *J. Phys. Chem.*, **100**, 16302 (1996).
10. S. M. Davis and G. A. Somorjai, *J. Phys. Chem.*, **87**, 1545 (1983); M. Salmerón and G. A. Somorjai, *J. Phys. Chem.*, **86**, 341 (1982); F. Zaera and G. A. Somorjai, *J. Phys. Chem.*, **89**, 3211 (1985).
11. F. Zaera and G. A. Somorjai, *J. Am. Chem. Soc.*, **106**, 2288 (1984).
12. P. S. Cremer, X. Su, Y. R. Shen, and G. A. Somorjai, *J. Am. Chem. Soc.*, **118**, 2942 (1996).
13. J. Mitchell, *J. Am. Chem. Soc.*, **80**, 5848 (1958).
14. H. Pines and C. T. Chen, *J. Org. Chem.*, **26**, 1057 (1961).
15. J. A. Feighan and B. H. Davis, *J. Catal.*, **4**, 594 (1965).